

## Pablo GAMALLO OTERO

CiTIUS (Centro Singular de Investigación en Tecnoloxías da Información) University of Santiago de Compostela  
Santiago de Compostela, Galiza  
phone : (+34) 981563100, ext. 11782 / email: pablo.gamallo@usc.es

Born 27 July 1969 in Vigo, Galiza, Spain

Current Positions:

- **Associate Professor**, University of Santiago de Compostela, Spain.
- Promoter and co-founder of **Cilenis**, a Spin-Off on language technologies.
- Member of [ProLNat@GE](mailto:ProLNat@GE), a research team on Natural Language Processing, and [CiTIUS](#)

---

### EDUCATION

- Mars 1998, **Ph.D in Linguistics**, Blaise Pascal University, France.  
(Grant supported by Galician Department of Trade and Industry, Spain)
- October 1993, **Master on Linguistics, Logic and Computing**, Blaise Pascal University, France.
- July 1992, **Graduated in Hispanic Languages**, University of Santiago de Compostela, Spain,  
(University of Vigo - Spain / University of Bourgogne – France)

---

### POSITIONS

- 2007 - 2013 **Ramón y Cajal Researcher**, University of Santiago de Compostela, Spain
- 2004 - 2007 **Parga Pondal Reseacher**, University of Santiago de Compostela, Spain
- 2002 - 2004 **Post-Doc** supported by Fundação da Ciência e a Tecnologia (FCT), Ref: SFRG / BDP / 1189 / 2002, center CITI, Faculdade de Ciência e Tecnologia, Universidade Nova de Lisboa, Portugal.
- 2000 - 2002 **Post-Doc** supported by Fundação da Ciência e a Tecnologia (FCT), PRAXIS XXI / BDP / 2213 / 99, center CENTRIA, Faculdade de Ciência e Tecnologia, Universidade Nova de Lisboa, Portugal.
- 1999 - 2000 **Auxiliar Professor** (Asociado P3) at University of Vigo, Spain
- 1998 -1999 **Auxiliar Professor** (ATER) at University of Blaise Pascal, France

---

### TEACHING (UNIVERSITY COURSES)

- 2012-17 **Tools for Natural Language Processing**  
Master course, University of Santiago de Compostela.
- 2012-17 **Spanish Language (Syntax and Lexicon)**  
Spanish Department, University of Santiago de Compostela.
- 2016-17 **Programming in Perl for Linguists** (10h/year)  
PhD program, University of Santiago de Compostela.
- 2009-12 **Methods in corpus linguistics and natural language processing** (15h/year)  
Master course, University of Santiago de Compostela, Spain.
- 2007-09 **Corpus linguistics: corpus elaboration and information extraction** (15h/year)  
PhD program, University of Santiago de Compostela.
- 2007-11 **Computational analysis of hispanic texts** (60h/year)  
Spanish Department, University of Santiago de Compostela.
- 2006-09 **Historical phonetics and phonology of Spanish** (30h/year)  
Spanish Department, University of Santiago de Compostela.
- 2004-09 **Spanish Language (phonetics, phonology, morphology, syntax)** (25h/year)  
Spanish Department, University of Santiago de Compostela
- 2003 **Introduction to Programming in Perl for Natural Language Processing**  
Master Course (15 hours), Faculty of Humanities, New University of Lisbon, Portugal
- 2001-02 **Semantic information extraction and thesaurus design**  
Master Course (11 hours), Faculty of Science and Technology, New University of Lisbon, Portugal
- 2001 **Generative Lexicon**

1999-00	PhD program (10 hours), University of Vigo. <b>Applied Linguistics</b> (60h) Linguistics and Translation Department, University of Vigo.
1997-99	<b>Computing and Language Acquisition</b> (50h/year) Linguistics Department, University of Blaise Pascal, France.
1998-99	<b>French Language (phonetics, phonology, morphology, syntax)</b> (30h) Linguistics Department, University of Blaise Pascal, France.
1998-99	<b>General Linguistics</b> (17h) Linguistics Department, University of Blaise Pascal, France.

---

#### RESEARCH

1993-99	<b>Member of "Laboratoire de Recherche sur le Langage" (LRL)</b> , Blaise Pascal University, France. Participation in 2 projects: ElaDyS (Elaboration Dynamique de la Signification), AMICAL (Architecture Multi-Agent Intelligente Compagnon d'Aide a l'Apprentissage de la Lecture).
2000-04	<b>Member of "Grupo de Língua Natural" (GLINt)</b> , Computing Science Department, New University of of Lisbon, Portugal. Participation in 2 projects: TRADAUT-PT (Machine Translation Portuguese-English, English-Portuguese, Portuguese-French, French-Portuguese), FASTLING (acción integrada Portugal-Francia).
since 2004	<b>Member of "Grupo Gramática do Espanhol" (GE)</b> , University of Santiago de Compostela, Spain. Participation in 2 projects: GARI-COTER (terminology extraction) and COLIBRI (question-answering), Ministerio de Educación y Ciencia.
since 2007	<b>Principal Investigator of the research line "Processamento da Língua Natural" (ProLNat@GE)</b> , funded by the Galician Government (2010-2013).
since 2012	<b>Member of Centro de Investigación en Tecnoloxías da Información (CiTIUS)</b> , University of Santiago de Compostela.

#### Projects as Principal Investigator:

<i>Title</i>	<b>Un método lingüístico-estadístico para la traducción automática basado en corpus no-paralelos y semántica composicional</b>
<i>Code</i>	-
<i>Duration</i>	01/10/2016 - 30/03/2018
<i>Amount</i>	39.974,91 €
<i>PI</i>	Pablo Gamallo Otero
<i>Funding</i>	Programa de Ayudas a Investigadores y Creadores Culturales Fundación BBVA
<i>Title</i>	<b>Tecnologías de la lengua para análisis de opiniones en redes sociales (Convenio con Univ. de Vigo)</b>
<i>Code</i>	FFI2014-51978-C2-1-R
<i>Duration</i>	06/10/2015 - 31/12/2017
<i>Amount</i>	3,000.00 €
<i>PI</i>	Pablo Gamallo Otero
<i>Funding</i>	MINECO – Convenio Univ. de Vigo / Univ. Santiago de Compostela
<i>Title</i>	<b>OntoPedia: Extracción automática de información ontológica y enciclopédica acerca de entidades con nombre</b>
<i>Code</i>	FFI2010-14986
<i>Duration</i>	01/01/2011 - 31/12/2013

<i>Amount</i>	70,180.00 €
<i>PI</i>	Pablo Gamallo Otero
<i>Funding</i>	Ministerio de Ciencia e Innovación.
<i>Title</i>	<b>EXTRA-LEX: Extracción automática de léxicos bilingües Galego-Español e actualización dos recursos lexicográficos de motores de tradución automática</b>
<i>Code</i>	PGIDIT07PXIB20401PR
<i>Duration</i>	01/01/2007-31/10/2010
<i>Amount</i>	54,050.00 €
<i>PI</i>	Pablo Gamallo Otero
<i>Funding</i>	Consellaría de Economía e Industria de la Xunta de Galicia.
<i>Title</i>	<b>Automatización da análise sintáctica</b>
<i>Code</i>	RYC-2007-00905
<i>Duration</i>	20/12/2007-19/12/2009
<i>Amount</i>	15,000.00 €
<i>PI</i>	Pablo Gamallo Otero
<i>Funding</i>	Ministerio de Educación y Ciencia
<i>Title</i>	<b>Automatic Design of a Proper Noun Ontology for a Question-Answering System</b>
<i>Code</i>	HP2007-0061
<i>Duration</i>	01/01/2008-31/03/2010
<i>Amount</i>	8,300.00 €
<i>PI</i>	Pablo Gamallo Otero, Paulo Quaresma (Universidade de Évora, Portugal)
<i>Funding</i>	Ministerio de Educación y Ciencia. (Program of “Acciones Integradas”)
<i>Title</i>	<b>Consolidación e estruturación de unidades de investigación competitivas: grupos emerxentes</b>
<i>Code</i>	2008/101
<i>Duration</i>	01/01/2008-15/11/2010
<i>Amount</i>	75,000.00 €
<i>PI</i>	Pablo Gamallo Otero
<i>Funding</i>	Consellaría de Educación e Ordenación Universitaria de la Xunta de Galicia

#### Research Contracts University - Enterprise

<i>Title</i>	<b>EixOpenTrad: tradución automática avanzada de código aberto para as linguas de Galiza e Portugal</b>
<i>Code</i>	2007/CG258
<i>Duration</i>	21/02/2007-30/09/2007
<i>Amount</i>	3,850.00 €
<i>PI</i>	Pablo Gamallo Otero

<i>Partners</i>	Factoría de Software e Multimedia, Universidade de Santiago de Compostela
<i>Title</i>	<b>GalinOpenTrad: Tradución automática avanzada de código aberto para a integración europea da lingua galega</b>
<i>Code</i>	2008/CG269
<i>Duration</i>	1/04/2008-30/09/2008
<i>Amount</i>	2,300.00 €
<i>PI</i>	Pablo Gamallo Otero
<i>Partners</i>	Factoría de Software e Multimedia, Universidade de Santiago de Compostela
<i>Title</i>	<b>RecursoOpenTrad: Recursos lingüístico-computacionais de tradución automática avanzada en código aberto para a integración europea da lingua galega</b>
<i>Code</i>	2009/CG174-1
<i>Duration</i>	16/04/2009-30/09/2009
<i>Amount</i>	6,095.00 €
<i>PI</i>	Pablo Gamallo Otero
<i>Partners</i>	Factoría de Software e Multimedia, Universidade de Santiago de Compostela
<i>Title</i>	<b>COATI: Pescuda avanzada multilingüe en blogs para a recuperación de opinións e tendencias para o ámbito empresarial e da administración pública</b>
<i>Code</i>	2010/CG051
<i>Duration</i>	18/01/2010-30/10/2010
<i>Amount</i>	5,653.00 €
<i>PI</i>	Pablo Gamallo Otero
<i>Partners</i>	Factoría de Software e Multimedia, Universidade de Santiago de Compostela
<i>Title</i>	<b>CORUXA Biomedical Text Mining: extractor e codificador automático de información médica relevante mediante o uso da enxearía lingüística en código aberto</b>
<i>Code</i>	2011/CG338
<i>Duration</i>	10/05/2011-30/09/2011
<i>Amount</i>	9.200,00 €
<i>PI</i>	Pablo Gamallo Otero
<i>Partners</i>	Factoría de Software e Multimedia, Universidade de Santiago de Compostela
<i>Title</i>	<b>CELTIC- Coñecemento Estratéxico Liderado por Tecnoloxías para a Intelixencia Competitiva (CDTI-Feder-Innterconecta)</b>
<i>Code</i>	2012-CE138
<i>Duration</i>	21/09/2012-31/12/2014
<i>Amount</i>	21.000,00 €
<i>PI</i>	Pablo Gamallo Otero
<i>Partners</i>	Factoría de Software e Multimedia, Universidade de Santiago de Compostela

<i>Title</i>	<b>PLASTIC - PeopLe As a Service soportado por las Tecnologías de la Información y la Comunicación (CDTI-Feder-Innterconecta)</b>
<i>Code</i>	2013-CE298
<i>Duration</i>	01/04/2013-31/12/2015
<i>Amount</i>	31.154,00 €
<i>PI</i>	Pablo Gamallo Otero
<i>Partners</i>	INDRA SISTEMAS, S.A., Universidade de Santiago de Compostela

---

## PUBLICATIONS

### BOOKS

- Chambreuil M., Ben Gharbia A., Bernigot C., Gamallo P. Panissod C., Reinberger ML. (1998) *Sémantiques*. Paris, Editions Hermès. 416 pages. ISBN: 2-86601-721-8.
- Gamallo P. (1998) *Construction conceptuelle des expressions complexes: le traitement de la combinaison nom-adjectif*, Thèse à la Carte, Editorial Septentrion, Lille, 420 pages. ISBN: 2-284-00938-7.

### BOOK CHAPTERS

- Gamallo, Pablo (2017) "Linguakit y sus aplicaciones en el aula". In Domínguez M.J and Sanmarco T. (Eds.), *Lexicografía y didáctica*. Studien zur romanischen Sprachwissenschaft und interkulturellen Kommunikation. Peter Lang Edition. (pp. 229-246). DOI: 10.3726/978-3-653-05627-3. ISBN: 978-631-66448-3. General Ranking SPI: 6.
- Garcia, Marcos and Pablo Gamallo (2015). "Yet another suite of multilingual NLP tools". In: José-Luis Sierra-Rodríguez, José-Paulo Leal, Alberto Simões, *Languages, Applications and Technologies, Volume 563 of the series Communications in Computer and Information Science*, Springer, pp 65-75. DOI: 10.1007/978-3-319-27653-3\_7. ISBN: 978-3-319-27652-6. General Ranking SPI: 4.
- Gamallo P. Garcia, Marcos, del Río, I., González, I. (2015) "Avalingua: Natural Language Processing for Automatic Error Detection". In: Marcus Callies, Sandra Götz (Eds.), *Learner Corpora in Language Testing and Assessment*. John Benjamins Publishing Company (35-58). DOI: 10.1075/scl.70.02gam. ISBN: 978-90-272-0378-6. General Ranking SPI: 14.
- Gamallo P. (2012) "Propuesta para una semántica de las dependencias sintácticas". In: Tomás J. Juliá, Belén López, Victoria Vázquez and Axendadre Veiga (Eds.), *Cum Corde Et In Nova Grammatica: Estudios ofrecidos a Guillermo Rojo*. Publicacións Universidade de Santiago de Compostela, (341-351). ISBN: 978-84-9887-914-8.
- Agustini, A., Gamallo P., Lopes, G.P. (2004) "Lexical Learning for Attachment Resolution". In: A. Branco, A. Mendes and R. Ribeiro (Eds.), *Language Technology for Portuguese: Shallow Processing Tools and Resources*. Edições Colibri, Portugal, (105-120). ISBN: 699-9674-6.
- Gamallo P. (2003) "Categorías morfosintácticas, relaciones sintácticas y composicionalidad semántica". In: Clara Molina and Manuela Romano (Eds.), *Cognitive Linguistics in Spain at the Turn of the Century* (173-188). ISBN: 699-9674-6.
- Gamallo P. (2001) "Caracterización Semántico-Cognitiva de las Categorías Gramaticales Fundamentales". In: Augusto Soares da Silva (Ed.), *Linguagem e Cognição*, Associação Portuguesa de Lingüística, Braga, Portugal, (355-374). ISBN: 972-98336-5-6.

### JOURNALS INDEXED IN SCI/SCOPUS

- Gamallo, Pablo, Iván Rodríguez-Torres, Marcos Garcia, (2017). "Distributional Semantics for Diachronic Search", *Computers and Electrical Engineering*, available on-line 26 July 2017. DOI: <https://doi.org/10.1016/j.compeleceng.2017.07.017>. ISSN: 0045-7906. Q3 (JCR), Q1 (SJR).
- Gamallo, Pablo, José Ramom Pichel, Iñaki Alegria, (2017). "From language identification to language distance", *Physica A*, Vol.

- 484, pp. 152-162. DOI: 10.1016/j.physa.2017.05.011. ISSN: 0378-4371. Q1 (JCR), Q1 (SJR).
- Pablo Gamallo (2017), "The role of syntactic dependencies in compositional distributional semantics" *Corpus Linguistics and Linguistic Theory*, Ahead of print (Jan 2017). DOI: 10.1515/cllt-2016-0038. ISSN: 1613-7027. Q2 (JCR), Q1 (SJR).
- Gamallo, Pablo and Marcos Garcia (2017) "LinguaKit: uma ferramenta multilingue para a análise linguística e a extração de informação", *Linguamática*, 9(1), pp. 19-28. DOI: 10.21814/lm.9.1.243. ISSN: 1647-0818. Q4 (SJR).
- Pablo Gamallo (2016), "Comparing explicit and predictive distributional semantic models endowed with syntactic contexts" *Language Resources and Evaluation*, First online: 13 May 2016. DOI: 10.1007/s10579-016-9357-4. ISSN: 1574-0218. Q4 (JCR), Q1 (SJR).
- Gamallo, P. (2016) "Entity Linking with Distributional Semantics". *Lecture Notes in Computer Science*, vol. 9727, (177-188). Springer-Verlag. DOI: 10.1007/978-3-319-41552-9\_18. ISSN: 0302-9743. Q2 (SJR).
- Zubiaga, Arkaitz, Iñaki San Vicente, Pablo Gamallo, José Ramón Pichel, Alegria, Iñaki, Nora Aranberri, Aitzo Ezeiza and Víctor Fresno (2016), TweetLID: a benchmark for tweet language identification, *Language Resources and Evaluation*, 50(4), pp. 729-766. DOI: 10.1007/s10579-015-9317-4. ISSN: 1574-0218. Q4 (JCR), Q1 (SJR).
- Garcia M., Gamallo P. (2015) "Exploring the Effectiveness of Linguistic Knowledge for Biographical Relation Extraction", *Natural Language Engineering*, 21(4), pp. 519-551. DOI: 10.1017/S1351324913000314. ISSN: 1351-3249. Q4 (JCR), Q1 (SJR).
- Alegria, Iñaki, Nora Aranberri, Víctor Fresno, Pablo Gamallo, Lluís Padró, Iñaki San Vicente, Jordi Turmo and Arkaitz Zubiaga (2015), TweetNorm: a benchmark for lexical normalization of Spanish tweets, *Language Resources and Evaluation*, vol 49(4), 883-905, DOI: 10.1007/s10579-015-9315-6. ISSN: 1574-0218. Q4 (JCR), Q1 (SJR).
- Gamallo, P., Garcia, M. (2015) "Multilingual Open Information Extraction". *Lecture Notes in Computer Science*, vol. 9273, (711-722). Springer-Verlag. ISSN: 0302-9743. DOI: 10.1007/978-3-319-23485-4\_72. Q2 (SJR).
- Gamallo P. (2014) "Uso de corpora comparáveis para filtrar dicionários bilíngues gerados por transitividade", *DELTA: Documentação de Estudos em Linguística Teórica e Aplicada*, 30(2), (213-235). DOI: 10.1590/0102-445034728151307539. ISSN: 0102-4450. Q2 (SJR).
- Gamallo, Pablo, Juan Carlos Pichel, Marcos Garcia, José Manuel Abuín and Tomás Fernández Pena, (2014) "Análisis morfosintáctico y clasificación de entidades nombradas en un entorno Big Data". *Procesamiento del Lenguaje Natural*, 53, p. 17-24. ISSN: 1135-5948. Q2 (SJR).
- Garcia, Marcos and Pablo Gamallo, (2014) "Entity-Centric Coreference Resolution of Person Entities for Open Information Extraction". *Procesamiento del Lenguaje Natural*, 53, p. 25-32. ISSN: 1695-2618. Q2 (SJR).
- Garcia, Marcos, Pablo Gamallo, Iria Gayo and Miguel Anxo Pousada Cruz, (2014). "PoS-tagging the Web in Portuguese. National varieties, text typologies and spelling systems". *Procesamiento del Lenguaje Natural*, 53, p. 95-101. ISSN: 1695-2618. Q2 (SJR).
- Gamallo P. (2013) "Lexical inheritance with meronymic relationships", *Axiomathes*, vol. 23(1), (165-185), Springer Science. DOI 10.1007/s10516-011-9152-1, ISSN: 1121-1151. Q2 (SJR).
- Saralegi, X., Gamallo, P. (2013) "Analyzing the Sense Distribution of Concordances Obtained by Web As Corpus Approach". *Lecture Notes in Computer Science*, vol. 7816, (355-367), Springer-Verlag. DOI: 10.1007/978-3-642-37247-6\_29. ISSN: 0302-9743. Q2 (SJR).
- Gamallo, P., Garcia, M. (2012) "Extraction of Bilingual Cognates from Wikipedia". *Lecture Notes in Computer Science*, vol. 7243, (63-72). Springer-Verlag. DOI:10.1007/978-3-642-28885-2\_7. ISSN: 0302-9743. Q2 (SJR).
- Gamallo P. , Bordag S. (2011) "Is Singular Value Decomposition Useful for Word Similarity Extraction?", *Language Resources and Evaluation*, 45(2), (95-119). DOI: 10.1007/s10579-010-9129-5. ISSN: 1574-020X. Q4 (JCR), Q2 (SJR).
- Gamallo P. , González I. (2011) "A Grammatical Formalism based on Patterns of Part-of-Speech Tags", *International Journal of Corpus Linguistics*, 16(1), 45-71. DOI: 10.1075/ijcl.16.1.03gam. ISSN: 1384-6655. Q2 (JCR), Q1 (SJR).
- Gamallo, Pablo, Marcos Garcia (2011) "A Resource-Based Method for Named Entity Extraction and Classification". *Lecture Notes in Computer Science*, vol. 7026, (610-623). Springer-Verlag. DOI: 10.1007/978-3-642-24769-9\_44. ISSN: 0302-9743. Q2 (SJR).
- Gamallo P., Pichel J-R. (2010) "Automatic Generation of Bilingual Dictionaries Using Intermediary Languages and Comparable Corpora", *Lecture Notes in Computer Science*, vol. 6008, Springer-Verlag, (473-483). DOI: 10.1007/978-3-642-12116-6\_40. ISSN: 0302-9743. Q2 (SJR).

- Gamallo P. (2009) "Comparing Different Properties Involved in Word Similarity Extraction", *Lecture Notes in Computer Science*, vol. 5816, Springer-Verlag, (634-645). DOI: 10.1007/978-3-642-04686-5\_52. ISSN: 0302-9743. Q2 (SJR).
- Gamallo P. (2008) "Comparing Window and Syntax Based Strategies for Semantic Extraction", *Lecture Notes in Computer Science*, vol. 5190, Springer-Verlag, (41-50). DOI: 10.1007/978-3-540-85980-2\_5. ISSN: 0302-9743. Q2 (SJR).
- Gamallo P., Lopes G.P., Agustini A. (2008) "Automatic Acquisition of Formal Concepts from Text", *Journal for Language Technology and Computational Linguistics* (former LDV-Forum), 23(1), (59-74). ISSN: 0175-1336.
- Gamallo P., Pichel, J.R. (2008) "Learning Spanish-Galician Translation Equivalents Using a Comparable Corpus and a Bilingual Dictionary", *Lecture Notes in Computer Science*, vol. 4919, Springer-Verlag, (423-433). DOI:10.1007/978-3-540-78135-6\_36. ISSN: 0302-9743. Q2 (SJR).
- Gamallo P., Lopes G.P., Agustini A. (2007) "Inducing Classes of Terms from Text", *Lecture Notes in Computer Science*, vol. 4629, Springer-Verlag, (31-38). DOI: 10.1007/978-3-540-74628-7\_7. ISSN: 0302-9743. Q2 (SJR).
- Gamallo P., (2006) "Using Natural Alignment to Extract Translation Equivalents", *Lecture Notes in Computer Science*, vol. 3960, Springer-Verlag, (41-49). DOI: 10.1007/2F11751984\_5 . ISSN: 0302-9743. Q2 (SJR).
- Gamallo P., Agustini A. Lopes G.P. (2005) "Clustering Syntactic Positions with Similar Semantic Requirements". *Journal of Computational Linguistics*, 31(1), MIT Press, (107-146). DOI:10.1162/0891201053630318. ISSN: 0891-2017. Q1 (JCR), Q1 (SJR).
- Gamallo P., Pichel, J.R. (2005) "An Approach to Acquire Word Translations from Non-Parallel Texts", 12th Portuguese Conference on Artificial Intelligence (EPIA'05), *Lecture Notes in Computer Science*, vol. 3808, Springer-Verlag, (600-610). DOI: 10.1007/11595014\_59. ISSN: 0302-9743 / ISBN 3-540-23410-1. Q4 (JCR), Q2 (SJR).
- Gamallo P., Da Silva J. Lopes G.P. (2004) "A Divide-And-Conquer Approach to Acquire Syntactic Categories", In: C. Bento, A. Cardoso, and G. Dias (Eds.), International Conference of Grammatical Inference, *Lecture Notes in Computer Science*, vol. 3264, (151-162), Springer-Verlag. DOI: 10.1007/2F978-3-540-30195-0\_14. ISSN: 0302-9743 / ISBN 3-540-30737-0. Q4 (JCR), Q2 (SJR).
- Kozareva Z., Da Silva J., Gamallo P., Lopes G.P., (2004) "Cluster Analysis of Named Entities", In: M. Klopotek (Eds.), International Intelligent Information Processing and Web Mining Conference, *Advances in Soft Computing*, Vol. XIV, Springer-Verlag, (429-433). DOI: 10.1007/978-3-540-39985-8\_47. ISSN: 1615-3871 / ISBN: 3-540-21331-7. Q3 (SJR).
- Gamallo P. (2003) "Cognitive characterisation of basic grammatical structures". *Pragmatics & Cognition*, 11(2), Jonh Benjamins Publishing Company, (209-240). DOI: 10.1075/pc.11.2.03ote ISSN: 0929-0907. (indexed since 2006 in JCR)
- Gamallo P., Agustini A. Lopes G.P. (2003) "Acquiring Semantic Classes to Elaborate Attachment Heuristics", In: F. Moura Pires and S. Abreu (Eds.), 11th Portuguese Conference on Artificial Intelligence (EPIA'03), *Lecture Notes in Computer Science*, Vol. 2902, Springer (479-488). DOI: 10.1007/2F978-3-540-24580-3\_56. ISSN: 0302-9743 / ISBN: 3-540-20589-6. Q2 (SJR).
- Agustini A., Gamallo P., Lopes G.P. (2003) "Selection Restrictions Acquisition for Parsing Improvement". In: O. Bartenstein, U. Geske, M. Hannebauer, and O. Yoshie (eds.), *Web-Knowledge Management and Decision Support* (Selected papers from the 14th International Conference on Applications of Prolog - INAP). *Lecture Notes in Computer Science*, Vol. 2543, Springer-Verlag, (129-146). DOI: 10.1007/3-540-36524-9\_11. ISSN: 0302-9743 / ISBN: 3-540-00680-X. Q2 (SJR).
- Agustini A., Gamallo P., Lopes G.P. (2002): "Assessment of Selection Restrictions Acquisition", In: G. Bettencourt and G. Ramalho (Eds.), 16th Brazilian Symposium on Artificial Intelligence, *Lecture Notes in Computer Science*, Vol. 2507, Springer (407-417). DOI: 10.1007/2F3-540-36127-8\_39. ISSN: 0302-9743 / ISBN: 3-540-00124-7. Q2 (SJR).
- Gamallo P., Agustini A. Lopes G.P. (2001) "Selections Restrictions Acquisition from Corpora", In: Pavel Brazdil and Alípio Jorge (Eds.), 10th Portuguese Conference on Artificial Intelligence (EPIA'01), *Lecture Notes in Computer Science*, Vol. 2258, Springer (30-43). DOI: 10.1007/2F3-540-45329-6\_7. ISSN: 0302-9743 / ISBN: 3-540-43030-X. Q2 (SJR).
- Gamallo P., Gasperin C. Agustini A. Lopes G.P. (2001) "Syntactic-Based Methods for Measuring Word Similarity", In: V. Matousek, P. Mautner, R. Moucek and K. Moucek (Eds.), *Text, Speech and Dialogue (TSD-2001)*. *Lecture Notes in Computer Science*, Vol. 2166, Springer (116-125). DOI: 10.1007/3-540-44805-5\_15. ISSN: 0302-9743 / ISBN: 3-540-42557-8. Q2 (SJR).

## OTHER JOURNALS

- Gamallo P., Garcia, M., González, I., Muñoz, M., Del Río, I. (2013) "Learning verb inflection using Cilenis conjugators ", *Eurocall Review* , 21(1): 12-19. ISSN: 1695-2618.
- Gamallo, P., Garcia, M. (2012) "Técnicas de procesamiento del lenguaje natural en la Recuperación de Información", *Novática*, vol. 219 , pages 42-47. ISSN: 0211-2124.
- Garcia, Marcos and Pablo Gamallo, 2011. Resolución de Correferencia de Nombres de Persona para Extracción de Información Biográfica. *Procesamiento del Lenguaje Natural*, 47, p. 47-55. ISSN: 1135-5948.
- García Marcos, Gamallo, P. (2010) "Análise Morfosintáctica para o Português Europeu e Galego: Problemas, Soluções e Avaliação", *Linguamática*, 2(2), 59-67. ISSN: 1647-0818.
- Malvar Paulo, Pichel J-R., Senra O., Gamallo, P., García B. (2010) "Vencendo a escassez de recursos computacionais. Carvalho: Tradutor Automático Estatístico Inglês-Galego a partir de corpus paralelo EuroParl Inglês-Português", *Linguamática*, 2(2), 31-38. ISSN: 1647-0818.
- Gamallo P., González, I. (2009) "Una gramática de dependencias basada en patrones de etiquetas" *Procesamiento del Lenguaje Natural*, 43, (315-324). ISSN: 1135-5948.
- Pichel, J.R, Malvar, P., Senra, O., Gamallo P., García, A. (2009) "Carvalho: English-Galician SMT system form EuroParl English-Portuguese parallel corpus" *Procesamiento del Lenguaje Natural*, 43, (379-381). ISSN: 1135-5948.
- Gamallo P. (2008) "The Meaning of Syntactic Dependencies", *Linguistik Online*, 35(3), (33-53). DOI: 10.13092/lo.35.522. ISSN: 1615-3014.
- Gamallo P. and J.R. Pichel (2007) "Un método de extracción de equivalentes de traducción a partir de un corpus comparable castellano-gallego", *Procesamiento del Lenguaje Natural*, 39, pp. 241-248. ISSN: 1135-5948
- Gamallo P., Lopes G.P., Agustini A. (2007), "Extraction of Lexico-Semantic Classes from Text", *Publication Series of the Institute of Cognitive Science (PICS)*, Vol. 1-2007 (39-48). ISSN: 1610-5389.
- Barcala Mario, Eva Domínguez Noya, Pablo Gamallo, Marisol López, Eduardo Moscoso, Guillermo Rojo, Paula Santalla, Susana Sotelo. (2007) "El Proyecto Gari-Coter en el Seno del Proyecto RICOTERM2", *Procesamiento del Lenguaje Natural*, 39, pp. 295-296. ISSN: 1135-5948.
- Gamallo P., Gasperin C. Agustini A. Lopes G.P, Lima, V. (2005) "Using Syntactic Methods to Learn Semantic Information", *Linguistica Computazionale*, vol XXII-XXIII, (201-228). DOI: 10.1400/18228. ISSN: 88-8147-413-1.
- Gamallo P., Sotelo, S. (2005) "El tratamiento de la polisemia en la extracción de léxicos bilingües a partir de corpora paralelos". *Procesamiento del Lenguaje Natural*, 35, (103-110). ISSN: 1135-5948.
- Noncheva V., Gamallo P., Agustini A. Lopes G.P. (2004) "A Stochastic Approach for Finding of Semantically Related Words". *Pliska Studia Mathematica Bulgarica*, 16, (171-182). ISSN: 0204-9805.
- Gamallo P., Lopes G.P., Agustini A. (2004) "The Role of Optional Co-Composition to Solve Lexical and Syntactic Ambiguity". *Procesamiento del Lenguaje Natural*, 33 (73-80). ISSN: 1135-5948.
- Gamallo P., Agustini A. Lopes G.P. (2003) "Learning Subcategorisation Information to Model Grammar with Co-Restrictions". *Traitement Automatique des Langues*, 44(1), (93-118). ISSN: 1248-9433. (indexed since 2012).
- Gamallo P., Agustini A. Lopes G.P. (2002) "Usando la co-composicionalidad en la adquisición de la subcategorización sintáctico-semántica". *Procesamiento del Lenguaje Natural*, 29 (35-44). ISSN: 1135-5948.
- Gamallo P. (2000) "Bases lexicales et systèmes d'héritage conduits par la relation de méronymie". *Revue Française de Linguistique Appliquée*, 5(2) (pp. 45-56). ISSN 1386-1204. (indexed since 2004).
- Gamallo P. (2000) "La métonymie dans le processus d'interprétation d'expressions complexes". *Revue de Sémantique et Pragmatique*, 7 (pp. 29-58). ISSN: 1285-4093.
- Gamallo P. (2000) "Bases léxicas organizadas mediante un sistema de herencia mereológica". *Procesamiento del Lenguaje Natural*, 26, (pp. 65-72). ISSN: 1135-5948.
- Gamallo P. & Reinberger M-L (1999) "Modelización del proceso de combinación de estructuras léxicas". *Procesamiento del Lenguaje Natural*, 25, (pp. 83-91). ISSN: 1135-5948.
- Gamallo P. & Chambreuil M (1998) "Una modelización del mecanismo dinámico de construcción de la significación de expresiones complejas". *Novática*, 135, (pp. 50-54). ISSN 211-2124.
- Gamallo P. & Chambreuil M. (1998) "Léxico generativo y mecanismos de control en el proceso de interpretación". *Procesamiento*



*del Lenguaje Natural*, 23, (pp. 54-60). ISSN: 1135-5948.

Chambreuil M., Ben Gharbia A. & Gamallo P. (1998) "Variations sur la compositionnalité montaguienne". *Traitement Automatique de la Langue*, 39(1), (pp. 35-65). ISSN: 1248-9433.

Gamallo P. (1995) "Léxico e inferencia : una semántica de acceso a la información". *Procesamiento del Lenguaje Natural*, 17, (195-209). ISSN: 1135-5948.

## PROCEEDINGS OF INTERNATIONAL CONFERENCES

Gamallo, Pablo, (2017). "Citius at SemEval-2017 Task 2: Cross-Lingual Similarity from Comparable Corpora and Dependency-Based Contexts". In *Proceedings of 11th International Workshop on Semantic Evaluation (SemEval-2017), at ACL 2017*, Vancouver, Canada: 226-229. ISBN 978-1-945626-00-5.

Gamallo, Pablo, (2017). "Sense Contextualization in a Dependency-Based Compositional Distributional Model". In *Proceedings of 2nd Workshop on Representation Learning for NLP (Rep4NLP-2017), at ACL 2017*, Vancouver, Canada: 1-9. ISBN 978-1-945626-62-3.

Garcia, Marcos and Pablo Gamallo, (2017). "A rule-based system for cross-lingual parsing of Romance languages with Universal Dependencies". In *Proceedings of Conference on Computational Natural Language Learning (CoNLL-2017), at ACL 2017*, pp. 274-282, Vancouver, Canada. ISBN 978-1-945626-54-8.

Gamallo, Pablo, Iván Rodríguez-Torres and Marcos Garcia (2017). "A Web Interface for Diachronic Semantic Search in Spanish". In *Proceedings of the Software Demonstrations at the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2017)*, Valencia: 45-48. ISBN 978-1-945626-36-4

Gamallo, Pablo, José Ramon Pichel and Iñaki Alegria (2017). "A Perplexity-Based Method for Similar Languages Discrimination". In *Proceedings of Fourth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2017) at the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2017)*, Valencia: 109-114. ISBN 978-1-945626-43-2

Gamallo, Pablo, Martín Pereira-Fariña (2017). "Compositional Semantics using Feature-Based Models from WordNet". In *Proceedings of Workshop on Sense, Concept and Entity Representations and their Applications at the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2017)*, Valencia: 1-11. ISBN 978-1-945626-50-0.

Almatarneh, Sattam, Pablo Gamallo (2017) "Automatic Construction of Domain-Specific Sentiment Lexicons for Polarity Classification", F. De la Prieta et al. (eds.), Trends in Cyber-Physical Multi-Agent Systems, *15th International Conference on Practical Applications of Agents and Multi-Agent Systems (PAAMS 2017)*, Advances in Intelligent Systems and Computing, vol 619, Springer. DOI 10.1007/978-3-319-61578-3\_17. ISBN: 978-3-319-61577-6.

Iñaki San Vicente, Iñaki Alegria, Nora Aranberri, Cristina España-Bonet, Pablo Gamallo, Hugo Gonçalo Oliveira, Eva Martínez, Antonio Toral, Arkaitz Zubiaga (2016) TweetMT: A parallel microblog corpus. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*. ISBN 978-2-9517408-9-1.

Gamallo, Pablo (2015). "Dependency Parsing with Compression Rules". *Proceedings of the 14th International Conference on Parsing Technologies*, pages 107–117, Bilbao, Spain; July 22–24. ISBN 978-1-941643-98-3.

Garcia, Marcos and Pablo Gamallo, (2015). "Yet another suite of multilingual NLP tools". In J-L. Sierra, J-P. Leal and A. Simões, *Proceedings of the Symposium on Languages, Applications and Technologies (SLATE 2015)*, Madrid, Spain: 81-90. ISBN 978-84-606-8762-7.

Garcia, Marcos and Pablo Gamallo, (2014). "An Entity-Centric Coreference Resolution System for Person Entities with Rich Linguistic Information". In *Proceedings of the 25th International Conference on Computational Linguistics (COLING 2014)*, Dublin: 171-175. ISBN 978-1-941643-26-6.

Abuín, J.M., Juan C. Pichel, Tomás F. Pena, Pablo Gamallo and Marcos García (2014) "PerlDooop: Efficient Execution of Perl Scripts on Hadoop Clusters", *IEEE Int. Conference on Big Data (IEEE Big Data)*. Washington D.C., USA, October 2014.

Gamallo, Pablo and Marcos Garcia, (2014). "Citius: A Naive-Bayes Strategy for Sentiment Analysis on English Tweets". In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, Dublin: 171-175. ISBN 978-1-941643-24-2.

Gamallo, Pablo (2014). "An Overview of Open Information Extraction". In *Proceedings of the 3rd Symposium on Languages, Applications and Technologies (SLATE-2014)*, Bragança, Portugal: 13-16. ISBN: 978-3-939897-68-2. DOI:

- Garcia, Marcos and Pablo Gamallo, (2014). "Multilingual corpora with coreferential annotation of person entities". *Proceedings of the 9th edition of the Language Resources and Evaluation Conference (LREC 2014)*, Reykjavik: 3229-3233. ISBN: 978-2-9517408-8-4.
- Alegria, Iñaki, Nora Aranberri, Pere Comas, Víctor Fresno, Pablo Gamallo, Lluís Padró, Iñaki San Vicente, Jordi Turmo and Arkaitz Zubiaga (2014), "TweetNorm\_es: an Annotated Corpus for Spanish Microtext Normalization". En *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, European Language Resources Association (ELRA), Reykjavik, Iceland, ISBN: 978-2-9517408-8-4.
- Gamallo P., Garcia, M., González, I., Muñoz, M., Del Río, I. (2013) "An evaluation of Avalingua based on learner corpora", *ICAME34 Workshop Learner Corpora and their Application in Language Testing and Assessment*, May 22, Santiago de Compostela, Spain: 52-53.
- Gamallo P., Garcia, M., Fernández-Lanza, S.. (2012) "Multilingual Open Information Extraction", *EACL 2012 ROBUST-UNSUP Workshop*, April 24, Avignon, France. ISBN 978-1-937284-19-0 .
- Gamallo P., González, I. (2012) "DepPattern: A Multilingual Dependency Parser", Demo Session of the *International Conference on Computational Processing of the Portuguese Language (PROPOR 2012)*, April 17-20, Coimbra, Portugal.
- Garcia, Marcos and Pablo Gamallo (2011). "A Weakly-Supervised Rule-Based Approach for Relation Extraction". In Jose A. Lozano, Jose A. Gámez and José A. Moreno Pérez (eds.), *Proceedings of the XIV Conference of the Spanish Association for Artificial Intelligence (CAEPIA 2011). Workshop on Knowledge Extraction and Exploitation from Semi-structures Online Sources (KEESOS)*. La Laguna, Spain.
- Garcia, Marcos and Pablo Gamallo (2011). "Dependency-Based Text Compression for Semantic Relation Extraction". In Preslav Nakov, Zornitsa Kozareva, Kuzman Ganchev and Jerry Hobbs (eds.), *Proceedings of the Workshop on Information Extraction and Knowledge Acquisition (IEKA 2011)* at 8th International Conference on Recent Advances in Natural Language Processing (RANLP 2011), Hissar, Bulgaria: 21-28.
- Garcia, Marcos and Pablo Gamallo (2011). "Evaluating Various Linguistic Features on Semantic Relation Extraction". In Galia Angelova, Kalina Bontcheva, Ruslan Mitkov and Nikolai Mikolov (eds.), *Proceedings of the 8th International Conference on Recent Advances in Natural Language Processing (RANLP 2011)*, Hissar, Bulgaria: 721-726.
- Garcia, Marcos and Pablo Gamallo, 2011. An Exploration of the Linguistic Knowledge for Semantic Relation Extraction in Spanish. In Patrick Saint-Dizier and Rutu Mehta-Melkar (eds.), *Proceedings of the Joint Workshop FAMILBR/KRAQ'11. Learning by Reading and its Applications in Intelligent Question-Answering at 22nd International Joint Conference on Artificial Intelligence (IJCAI'11)*, Barcelona: 7-12.
- Gamallo P., González I. (2010) "Wikipedia as Multilingual Source of Comparable Corpora", *LREC Workshop on Building and Using Comparable Corpora*, May 17, Malta. ISBN: 2-9517408-6-7.
- González, I., Gamallo P.(2010) "La Wikipedia como fuente multilingüe de corpus comparables", *II Congreso Internacional de lingüística de Corpus (CILC-2010)*, May 13-15, A Coruña, pp. 369-378. ISBN: 978-84-9749-401-4 .
- Malvar, P., Pichel, J.R., Senra, O., Gamallo P., García, A. (2010) "Obtaining computational resources for languages with scarce resources from closely related computationally-developed languages. The Galician and Portuguese case ", *II Congreso Internacional de lingüística de Corpus (CILC-2010)*, May 13-15, A Coruña., pp. 529-536. ISBN: 978-84-9749-401-4 .
- García, M., Gamallo P.(2010) "Using morphosyntactic post-processing to improve PoS-tagging accuracy", *International Conference of Computational Processing of Portuguese Language (PROPOR 2010)*, Porto-Alegre, Brasil. ISSN: 2177-3580
- Gamallo P. (2008) "Evaluating two different methods for the task of extracting bilingual lexicons from comparable corpora", In Proceedings of *LREC Workshop on Comparable Corpora*, Marrakech, Marroco, pp. 19-26. ISBN: 2-9517408-4-0.
- Gamallo P. (2007) "Learning Bilingual Lexicons from Comparable English and Spanish Corpora", In Proceedings of *Machine Translation Summit XI*, Copenhagen, Denmark, pp. 191-198. ISBN: 978-87-90708-16-0.
- Barcala Mario, Eva Domínguez Noya, Pablo Gamallo, Marisol López, Eduardo Moscoso, Guillermo Rojo, Paula Santalla, Susana Sotelo. (2007) "A Corpus and Lexical Resources for Multi-word Terminology Extraction in the Field of Economy", *3rd Language & Technology Conference (LeTC'2007)*, Poznan, Poland (355-359). ISBN: 978-83-7177-407-2.
- Gamallo P., Lopes G.P., Agustini A. (2006) "Extraction of Lexico-Semantic Classes from Text", *International Workshop on Ontologies in Text Technology (OTT'06)*, Osnabrück, Germany (39-44).

- Gamallo P., Da Silva, J., Lopes, G.P. (2005) "Cross-Lingual Classification of Function Words", In: Alexis Quesada, Roberto Moreno, José Carlos Rodríguez (Eds.), *10th International Conference on Computer Aided Systems Theory, Eurocast'05*, Las Palmas, Spain, (92-95). ISBN: 84-689-0432-5.
- Gamallo P. (2005) "Extraction of Translation Equivalents from Parallel Corpora Using Sense-Sensitive Contexts", In: J. Hutchins, B. Kis and G. Prószyński (Eds.), *10th Conference of the European Association for Machine Translation (EAMT'05)*, Budapest, Hungary (97-102). ISBN: 963-9206-04-0.
- Noncheva V. Gamallo P., Agustini A., (2003) "Automatic acquisition of Word Selection Restrictions: a Stochastic Approach". In: Ruslan Mitkov (ed.), *International Conference of Recent Advances in Natural Language Processing (RANLP'03)*. Borovets, Bulgaria, (347-351). ISBN: 954-90906-6-3.
- Noncheva V. Gamallo P., Agustini A., (2003) "A stochastic approach for finding of semantically related words". Tenth *International Summer Conference on Probability and Statistics (Seminar on Statistical Data Analysis - SDA'2003)*. Sozopol, Bulgaria (26-28).
- Gamallo P., Gonzalez, M., Agustini A. Lopes G.P. de Lima, V. (2002) "Mapping Syntactic Dependencies into Semantic Relations", *European Conference of Artificial Intelligence (ECAI'02), Workshop Natural Language Processing and Machine Learning for Ontology Engineering*, Lyon, France, (15-22).
- Gamallo P., Agustini A. Lopes G.P. (2002) "Using Co-composition for Acquiring Syntactic and Semantic Subcategorisation", *ACL'02 Workshop on Unsupervised Lexical Acquisition*, Philadelphia. Proceedings Published by ACL Office. (34-41).
- Gasperin C. Gamallo P., Agustini A. Lopes G.P. Lima V.L (2001) "The use of syntactic context for measuring word similarity", *ESSLLI-2001, Workshop on Semantic Knowledge Acquisition and Categorisation*, Helsinki, Finland.
- Agustini A., Gamallo P., Lopes G.P. (2001) "Selection Restrictions Acquisition for Parsing and Information Retrieval Improvement". *14th International Conference on Applications of Prolog (INAP'01)*, University of Tokyo, Tokyo, Japan (466-475). ISSN 1345-0980
- Gamallo P. (2000) "Lexical Inheritance in Upper-Level Ontologies, In: Kiril Simov and Atanas Kiryakov (Eds.), *Workshop on Ontologies and Lexical Knowledge Bases (OntoLex2000)*, Sozopol, Bulgaria (pp. 200-214).
- Gamallo P. & Chambreuil M. (1996) "Building up the meaning of problematic "verb+complements" constructions: the Co-Specification Device". *International Workshop of Predicative Forms in Natural Language and in Lexical Knowledge Bases*, Toulouse, France (pp. 89-98).

## PROCEEDINGS OF LOWER IMPACT CONFERENCES

- Martínez-Castaño, R., J.C. Pichel and P. Gamallo (2016) "Sentiment Analysis on Multilingual Tweets using Big Data Technologies". In proceedings of XXVII Jornadas de Paralelismo, Salamanca (España), pp. 119-126. ISBN: 978-84-9012-626-4.
- Alegria, Iñaki, Nora Aranberri, Cristina España-Bonet, Pablo Gamallo, Hugo Gonçalo Oliveira, Eva Martínez Garcia, Iñaki San Vicente, Antonio Toral, Arkaitz Zubiaga (2015), "Overview of TweetMT: A Shared Task on Machine Translation of Tweets at SEPLN 2015". En *Proceedings of the Tweet Translation Workshop 2015 co-located with 31st Conference of the Spanish Society for Natural Language Processing (SEPLN 2015)*, Alicante, Spain, CEUR Proceedings, pp. 8-19. ISSN: 1613-0073.
- Gamallo P, Garcia, M. Sotelo, S., and Pichel, J.R. (2014) "Comparing Ranking-based and Naive Bayes Approaches to Language Detection on Tweets ". In proceedings of *XXX Congreso de la Sociedad Española de Procesamiento de lenguaje natural. TweetLID: Twitter Language Identification Workshop at SEPLN 2014*, Girona, Spain. Spain. CEUR Proceedings, pp. 12-16. ISSN 1613-0073.
- Zubiaga, Arkaitz, Iñaki San Vicente, Pablo Gamallo, J.R. Pichel, Alegria, Iñaki, Nora Aranberri, Aitzol Ezeiza and Víctor Fresno (2014) "Overview of TweetLID: Tweet Language Identification at SEPLN 2014 ". In proceedings of *XXX Congreso de la Sociedad Española de Procesamiento de lenguaje natural. TweetLID: Twitter Language Identification Workshop at SEPLN 2014*, Girona, Spain. CEUR Proceedings, pp. 1-11. ISSN 1613-0073.
- Gamallo P, Garcia, M. and Fernández-Lanza, S. (2013) "TASS: A Naive-Bayes strategy for sentiment analysis on Spanish tweets". In proceedings of *XXIX Congreso de la Sociedad Española de Procesamiento de lenguaje natural. Workshop on Sentiment Analysis at SEPLN (TASS2013)*. Madrid. pp. 126-132. ISBN: 978-84-695-8349-4.
- Gamallo P, Garcia, M. and Pichel, J.R. (2013) "A Method to Lexical Normalisation of Tweets". In proceedings of *XXIX Congreso*

de la Sociedad Española de Procesamiento de lenguaje natural. *Workshop on Tweet Normalization at SEPLN*. Madrid. pp. 81-85. ISBN: 978-84-695-8349-4.

- Alegria, Iñaki, Nora Aranberri, Víctor Fresno, Pablo Gamallo, Lluís Padró, Iñaki San Vicente, Jordi Turmo and Arkaitz Zubiaga (2013) "Introducción a la tarea compartida Tweet-Norm 2013 : Normalización léxica de tuits en español". In proceedings of *XXIX Congreso de la Sociedad Española de Procesamiento de lenguaje natural. Workshop on Tweet Normalization at SEPLN*. Madrid. pp. 38-46. ISBN: 978-84-695-8349-4.
- Gamallo P, González, I. (2011) "Measuring Comparability of Multilingual Corpora Extracted from Wikipedia", *Workshop ICL on Iberian Cross-Language NLP tasks.*, Huelva (España), pp. 1-9. ISSN:1613-0073.
- García, M., Gamallo P.(2010) "Do preprocesamento morfológico à análise sintáctica de corpora multilíngue", *XXXIX Simposio Internacional de la Sociedad Española de Lingüística*, Santiago de Compostela, February 1-4. ISBN: 978-84-693-8655-2
- González, I., Gamallo P.(2010) "Estrategias para la elaboración de corpus comparables a partir de la web", *XXXIX Simposio Internacional de la Sociedad Española de Lingüística*, Santiago de Compostela, February 1-4. ISBN: 978-84-693-8655-2
- Gamallo P., Agustini A., Lopes P.G. (2004) "Disambiguation and Optional Co-composition". *Traitement Automatique de la Langue Naturelle (TALN'04)*, Fès, Marroco (199-204). ISBN: 2-9518233-4-7.
- Agustini A., Gamallo P., Lopes G.P. (2003): "Lexical Learning for Attachment Resolution", In: Anónio Branco (Ed.), *Workshop on Tagging and Shallow Processing of Portuguese (TASHA'03)*, Lisboa, Portugal (1-4).
- Gamallo P., Quaresma, P., Agustini A. Lopes G.P. (2002) "Using semantic word classes in text information retrieval systems". In Renata V. (ed.), *SBIE'02 - XII Symposium Brasileiro de Informática na Educação, Workshop de Ontologias*, Porto Alegre, Brazil (593-597). ISBN: 85-7431-133-2.
- Gamallo P., Agustini A., Lopes G.P. (2001) "The role of co-specification for the acquisition of selection restrictions from unsupervised corpora". *AFIA-2001, Workshop Applications Apprentissage, Acquisition des connaissances à partir de Textes Electroniques (A3CTE)*, Grenoble, France (27-33).
- Gamallo P. & Reinberger M-L (1999) "Activation de l'information lexicale dans la combinaison nom-adjectif". *Traitement Automatique de la Langue Naturelle (TALN99), Workshop Description des Adjectifs pour les Traitements Informatiques*, Corse, France (79-88).
- Gamallo P. & Chambreuil M (1996) "Une approche non modulaire de la Sémantique Lexicale". *Journées de Sémantique Lexicale Brestoises (JSLB-96)*, Brest, France.

---

## COMMITTEES

### Scientific Committees of Journals and Conferences:

- ◆ Revista Linguamática (desde 2010).
- ◆ Revista Agalia (desde 2012).
- ◆ SLATE 2012, 2013, 2014, 2015, 2017. Symposium on Languages, Applications and Technologies.
- ◆ EPIA 2001, 2005, 2007, 2009, 2011, 2013, 2015, 2017. Encontro Português de Inteligência Artificial.
- ◆ EMLex 2017, 4th Colloquium on Lexicography.
- ◆ LREC 2004, 2014, 2016 Conference on Language Resources and Evaluation.
- ◆ PROPOR 2006, 2008, 2010, 2012, 2014, 2016, Conference on Computational Processing of the Portuguese Language.
- ◆ CERI 2010, 2012, 2014, 2016 Conferencia Española de Recuperación de Información.
- ◆ PaCor 2016, Parallel Corpora: Creation and Applications International Symposium
- ◆ STILL 2009, 2011. Brazilian Symposium in Information and Human Language Technology.
- ◆ VERB 2010: Interdisciplinary Workshop on Verbs. The Identification and Representation.
- ◆ ISDA 2008. 2nd Workshop on Intelligent Text Categorization and Clustering (WITCC 2008).
- ◆ ACL 2005 Student Workshop. Michigan, EEUU.
- ◆ TEMA 2005, Workshop on Text Mining and Applications. Covilhã, Portugal.

### Organizing Committees:

- ◆ EMLex 2017, 4th Colloquium on Lexicography.
- ◆ PROPOR-2016 Co-located Workshops Chair.
- ◆ Twitter Machine Translation Workshop at SEPLN, TweetsMT, 2015.
- ◆ Twitter Language Identification Workshop at SEPLN, TweetsLID, 2014.
- ◆ Twitter Language Normalization Workshop at SEPLN, TweetNorm, 2013.

### Reviewer in international Journals and Conferences:

- ◆ Applied Sciences (2017)
- ◆ Information Science (2017)
- ◆ Algorithms (2017)
- ◆ Evolving Systems (2017)
- ◆ SemEval, Multilingual and Cross-lingual Semantic Word Similarity (2017)
- ◆ VarDial, Workshop on NLP for Similar Languages, Varieties and Dialects, co-located with EACL 2017.
- ◆ Revista Signos (2016)
- ◆ ComSIS journal (2015)
- ◆ Natural Language Engineering (2015)
- ◆ Traitement Automatique de la Langue (2013, 2014)
- ◆ SemEval, International Workshop on Semantic Evaluation (2014)
- ◆ IEICE Transactions (2012)
- ◆ International Journal of Corpus Linguistics (2012)

---

## PhD SUPERVISOR

Alexandre Agustini: “Aquisição Automática de Subcategorização Sintático-Semântica e sua Utilização em Sistemas de Processamento da Lengua Natural”. PhD in Computer Science. Faculty of Science and Technology, New University of Lisbon. Thesis defense: November 2006.

M. Pilar Valverde Ibáñez: “Descripción cuantitativa del orden de las funciones clausales argumentales en español”. PhD in Linguistics. Faculty of Philology, University of Santiago de Compostela. Thesis defense: May 2009.

Marcos García González: “Extracção de Relações Semânticas. Recursos, Ferramentas e Estratégias”. PhD in Linguística. CITIUS, University of Santiago de Compostela. Thesis defense: December 2014. Premio Extraordinario de Doctorado (USC) and Best PhD Dissertation Award at PROPOR 2016.

---

## SEMINARS and TALKS (invited)

- **Distributional Semantics and Compositionality and Open Source Modules for NLP (Linguakit)**, Research Seminar, Computing Centre / Argumentation Technology Research Group, University of Dundee, Scotland (UK) 15/06/2017.
- **Distributional Semantics and Compositional Translation**, Research Insider, CITIUS, Univ. of Santiago de Compostela, 24/03/2017.
- **Strategies to Build High Quality Bilingual Lexicons from Comparable Corpora**, Parallel Corpora: Creation and Applications International Symposium PaCor-2016, Univ. of Santiago de Compostela, 01/12/2016.
- **Strategies for Open Information Extraction**, Keynote at LexSem+Logics 2016 Workshop, co-located in PROPOR-2016, Tomar, Portugal, July 2016.
- **Relaciones entre ciencia y empresa: situación y perspectivas de futuro**, Mesa Redonda: 25 Aniversario del Manifiesto Cotec de el Escorial, Fundación COTEC para la Innovación, Universidade de Santiago, December 2015.
- **Avalingua: Corrector y evaluador de la calidad lingüística de textos**, VII Jornadas Empresa-Universidad RedPliR, Universidade de Santiago de Compostela, November 2015.

- **Ferramentas Lingüísticas na USC**, Xornadas de Lexicografía, Faculdade de Filologia, Universidade de Santiago de Compostela, June 2015.
- **An Overview of Open Information Extraction and Linguakit**, Seminar at INESC, University of Porto, December 2014.
- **An Overview of Open Information Extraction**, Keynote at 3rd Symposium on Languages, Applications, and Technology (SLATE-2014), Instituto Politécnico de Bragança, June 2014.
- **Web Inteligente**, Mesa Redonda in summer course: Big Data & Data Science, July, 2013.
- **A Depurative Strategy for Dependency Parsing with Finite-State Transducers**, Seminario Parsing de Dependencias, Faculdade de Informática, Universidade da Coruña, June 2012.
- **Construção de dicionários bilingues por transitividade**, I Workshop Per-Fide, Construção, Exploração e aplicação de Corpora Paralelos, Universidade do Minho, September 2010.
- **Extração automática de thesaurus**, Jornadas de Informática, Universidade do Minho, September 2010.
- **Modelos lingüísticos para a educação**, II Jornadas da Língua, Universidade de Ourense, Janeiro 2010, e I Jornadas de Cultura, Língua e Ensino, Universidade da Coruña, Março, 2010.
- **Software libre na USC para o procesamento da linguaxe natural**, Summer courses of USC intituled: “O Software libre e a Lingüística”, september 2009.
- **Extracción de classes semânticas em Galois Lattices e extracción de léxicos bilingues a partir de corpora comparáveis não-paralelos**, Seminars of Centro de Lingüística, Faculdade de Letras da Universidade de Porto (FLUP), Mars, 2007.
- **Lingüística de corpus y extracción de información**, II Jornadas de Actualización Gramatical. Universidade de Santiago de Compostela, October, 2006 .
- **Extraction methods of bilingual lexicon from parallel and non-parallel corpora**, 1st International Workshop of Researchers, Universidade de Vigo, Mars 2006.
- **Cómo usar un corpus para identificar esquemas sintáctico-semánticos?**, Seminar SERES, Universidade de Vigo, 2005.
- **Thesaurus Design from Analised Corpora**, Seminars of *GLINt (Grupo de Lingua Natural)*, Universidade Nova de Lisboa, Portugal, Mars 2003.
- **A method for acquiring selection restrictions**, Seminars of *CNTS (Centrum voor Nederlandse Taal en Spraak)*, University of Antwerp, Belgium, July 2002.
- **Preliminary results on selection restrictions extraction from partially parsed texts collections**, *Lisbon Meeting of the TRADAUT-PT MLIS project*, Universidade Nova de Lisboa, Portugal, January 2001.
- **Um sistema de Pesquisa de Informação para Textos em Língua Portuguesa**, Tutorial: *Presente e Futuro do E-Learning*, Universidade Aberta, Lisboa, Portugal, June 2001
- **Semântica Lexical**, seminars of *INESC*, Lisbon, Portugal, April 2001.
- **Construction dynamique de la signification**, seminars of Laboratoire d'Informatique de Paris-Nord (LIPN). 1998.

---

## AWARDS

- 1º Premio in *XI Concurso de Proyectos Innovadores de la Universidade de Santiago de Compostela* (2012)
- Honorable Mention in *Building Global Innovators 2013*, MIT Portugal, Lisboa. (2013)